

# STORAGE SWITZERLAND

## CAN TAPE SAVE CLOUD STORAGE



George Crump, Lead Analyst

“Cloud Storage” - the term invokes a vision of an infinitely scalable disk storage system assembled from commodity hardware that can provide near-instant access to the data it stores. There are two problems with this concept of cloud storage ‘nirvana’. First is the inability to bulk load or unload data because of bandwidth constraints and second is the long term cost to store all this data, since it may be there for decades.

### The Disk Based Cloud Architecture

Moving most data assets to the cloud is in many ways a massive consolidation exercise. This exercise can involve millions of users sending files from their home computers or an enterprise trying to replace dozens of NAS systems. The ability to build very large, efficient, scalable storage architectures is important, but doing so at an affordable cost over the long term is the key.

In either the personal or corporate consolidation scenario, traditional file systems will not scale efficiently to meet the demand. A change is needed to provide a storage architecture that can scale while still utilizing a

high percentage of available capacity. Object storage is offered as the method to achieve those goals.

As a result most storage infrastructures (both private and public) that wear the "cloud" moniker use object based storage systems. These systems store data as discrete objects, which for the most part can be thought of as files. The object storage architecture gives each file a unique identifier similar to a serial number or thumbprint and allows data locations to be stored in a flat index. This flat-index approach requires much less metadata (the information required to store and handle data objects), making it more scalable than the traditional POSIX file system, that uses a more complex folder structure to organize data.

Thanks to this flat index, object storage systems can be run much closer to full capacity without performance impact and can store a virtually unlimited amount of objects. The interface to these systems is typically through a REST (REpresentational State Transfer) API using simple "get" and "put" commands to access data.

Object storage overcomes many of the problems of traditional file storage. We are even seeing the emergence of a standards approach to the way object storage is being interfaced with, like Simple Storage Services (S3), OpenStack Object Storage (SWIFT) and the Cloud Data Management Interface (CDMI). But there are still some key weaknesses to be addressed in terms of bulk data movement and the cost of long term retention.

#### The Disk Based Cloud Problem

One of the challenges that a user of cloud storage faces is the bandwidth limitation of their cloud connection. This is typically an issue during the initial 'seeding' process when very large data sets are sent to the cloud, however it can also occur when a lot of data needs to be pulled from the cloud repository as well. A restore operation would be the most common example. Another use case is when a cloud provider shuts down, as [Nirvanix](#) has recently, giving users a short period of time to get their data out of the provider's facility.

A second challenge is the long term cost of storing all this data, especially given the design of object storage, which is made up of a cluster of servers called "nodes". Each of these nodes has internal hard drives that are aggregated into a single pool of storage. While this can provide nearly limitless scalability and linear performance it also requires that all the nodes in the architecture be powered on. The acquisition cost of each node can seem like an insignificant, one-time investment, but the cost to power and cool those nodes is a recurring monthly expense that will never end.

Over years and potentially decades of use this cost can be overwhelming as architectures begin to reach multiple hundreds or even thousands of nodes. There is also a space issue. Each node consumes data center floor space and, as the expansion continues, can force the

organization to build a new data center or at least redesign the old one.

#### The Tape Fix

From the beginning of its existence the designers of cloud storage infrastructures have overlooked tape storage as a potential component in their storage architectures. These object storage systems were proudly disk based. Now may be the time to reconsider those decisions. If tape can be integrated into a cloud storage infrastructure it can specifically address the challenges of bulk data loading and long term storage costs.

Tape can be an ideal solution to both of these obstacles on the road to greater cloud storage adoption. First, tape is unmatched for moving large amounts of data into and out of the cloud. Tape is designed to store TBs of information per cartridge, it's ruggedized for transport and can deliver a larger payload to a remote site in a 24 hour period than can the fastest WAN.

Tape is also a near-line storage mechanism when used in conjunction with an automated tape library. Tape media can store data sitting on a shelf inside the library without consuming any power, until it's placed in a tape drive. Because of this, tape libraries like [Spectra Logic's T-Finity](#), can densely pack high capacity tape cartridges within their confines to deliver many more PBs per floor tile than disk arrays.

Using tape can greatly reduce power, cooling and floorspace costs as well. And, tape in the form of [LTO5 and LTO6](#) has a capability called the Linear Tape File System (LTFS), an open tape file system standard that allows for long-term interchange between applications, tape hardware and manufacturers. If a cloud storage provider could leverage tape they would have a significant advantage in terms of bulk data loads and long-term storage costs.

## Making the Tape-to-Cloud Connection

There are some challenges facing providers that would like to leverage tape's advantages. First is the concern over access performance. Users of cloud based storage systems, both public and private, expect NAS-like response times to data requests. While a few seconds may be acceptable a few minutes is typically not.

This access problem can be dealt with by carefully integrating tape with disk storage or even flash storage. There are plenty examples of this in the "Tape NAS" market. These solutions turn a tape library into a network mount point that's accessible by CIFS (SMB) or NFS. The tape library then leverages some sort of disk cache area so that active data can be served instantly. Essentially active data and new data are stored on the disk area first so that users see rapid response times.

But similar to the way traditional disk NAS has been replaced by object based solutions, the tape NAS concept needs to evolve so it can compliment disk based cloud storage. In order to support this new 'cloud reality' it needs to become object based. Instead of turning a tape library into a network mount point it needs to turn the tape library into an extension of the object store itself. This could be done by building a RESTful interface into the tape offering either through a gateway or natively within the library as Spectra Logic has done with their [DS3 interface](#).

A cloud storage system that leverages tape could also take advantage of LTO's LTF5 format. This allows for long-term readability by eliminating proprietary tape formats. It's also an ideal way to bulk transfer data between customer sites and even between clouds. Imagine if a provider had to shut down for some reason. With this scenario of tape in the cloud, all they would have to do is ship tapes to a provider of the customer's choice and they could be read directly into the new object storage environment.

## Conclusion

Tape, while largely ignored by cloud storage operators, is actually an ideal match for this business model. It allows for the cost effective storing and transferring of high quantities of data. These attributes of tape as well as data interchangeability are available now. The next step is for tape library vendors to create an object storage interface for their solutions that will allow the merging of these two worlds.

### **About Storage Switzerland**

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: <http://www.storage-switzerland.com>

*Copyright © 2013 Storage Switzerland, Inc. - All rights reserved*